OIPE

RAW SEQUENCE LISTING   DATE: 02/12/2002
PATENT APPLICATION: US/09/813,271B   TIME: 08:59:40

Input Set : **N:\Crf3\02062002\I813271B.raw**
Output Set: **N:\CRF3\02122002\I813271B.raw**

SEQUENCE LISTING
```
 3 (1) GENERAL INFORMATION:
 4      (i) APPLICANT: Nico Cerletti
 5     (ii) TITLE OF INVENTION: New process for the production of
 6                             biologically active protein
 7    (iii) NUMBER OF SEQUENCES: 13
 8     (iv) CORRESPONDENCE ADDRESS:
 9          (A) ADDRESSEE: Novartis Patent Department
10          (B) STREET: 564 Morris Avenue
11          (C) CITY: Summit
12          (D) STATE: New Jersey
13          (E) COUNTRY: USA
14          (F) ZIP: 07901
15      (v) COMPUTER READABLE FORM:
16          (A) MEDIUM TYPE: Floppy disk
17          (B) COMPUTER: IBM PC compatible
18          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
19          (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
20     (vi) CURRENT APPLICATION DATA:
C--> 21          (A) APPLICATION NUMBER: US/09/813,271B
C--> 22          (B) FILING DATE: 20-Mar-2001
23    (vii) PRIOR APPLICATION DATA:
24          (A) APPLICATION NUMBER: PCT/EP95/02719
25          (B) FILING DATE: 12-Jul-95
26          (A) APPLICATION NUMBER: EPO 94810439.3
27          (B) FILING DATE: 25-Jul-94
28   (viii) ATTORNEY/AGENT INFORMATION:
29          (A) NAME: Pfeiffer, Hesna J. .
30          (B) REGISTRATION NUMBER: 22640
31          (C) REFERENCE/DOCKET NUMBER: 4-20039C/ClCl/USN
32     (ix) TELECOMMUNICATION INFORMATION:
33          (A) TELEPHONE: (908) 522-6940
34          (B) TELEFAX: (908) 522-6955
35 (2) INFORMATION FOR SEQ ID NO: 1:
36      (i) SEQUENCE CHARACTERISTICS:
37          (A) LENGTH: 339 base pairs
38          (B) TYPE: nucleic acid
39          (C) STRANDEDNESS: double
40          (D) TOPOLOGY: linear
41     (ii) MOLECULE TYPE: cDNA to mRNA
42    (iii) HYPOTHETICAL: NO
43    (vii) IMMEDIATE SOURCE:
44          (B) CLONE: E. coli LC137/pPLMu.hTGF-beta1 (DSM 5656)
```

```
            RAW SEQUENCE LISTING              DATE: 02/12/2002
            PATENT APPLICATION:  US/09/813,271B      TIME: 08:59:40

            Input Set : N:\Crf3\02062002\I813271B.raw
            Output Set: N:\CRF3\02122002\I813271B.raw
```

```
45      (ix) FEATURE:
46           (A) NAME/KEY: CDS
47           (B) LOCATION:1..336
48           (D) OTHER INFORMATION:/product= "human TGF-beta1"
49      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
50        GCC CTG GAC ACC AAC TAT TGC TTC AGC TCC ACG GAG AAG AAC TGC TGC      48
51        Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys
52         1               5                  10                  15
53        GTG CGG CAG CTG TAC ATT GAC TTC CGC AAG GAC CTC GGC TGG AAG TGG      96
54        Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp
55                       20                  25                  30
56        ATC CAC GAG CCC AAG GGC TAC CAT GCC AAC TTC TGC CTC GGG CCC TGC     144
57        Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly Pro Cys
58                 35                  40                  45
59        CCC TAC ATT TGG AGC CTG GAC ACG CAG TAC AGC AAG GTC CTG GCC CTG     192
60        Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu Ala Leu
61            50                  55                  60
62        TAC AAC CAG CAT AAC CCG GGC GCC TCG GCG GCG CCG TGC TGC GTG CCG     240
63        Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys Val Pro
64         65                  70                  75                  80
65        CAG GCG CTG GAG CCG CTG CCC ATC GTG TAC TAC GTG GGC CGC AAG CCC     288
66        Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg Lys Pro
67                       85                  90                  95
68        AAG GTG GAG CAG CTG TCC AAC ATG ATC GTG CGC TCC TGC AAG TGC AGC     336
69        Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys Cys Ser
70                 100                 105                 110
71        TGA                                                                 339
73 (2) INFORMATION FOR SEQ ID NO: 2:
74      (i) SEQUENCE CHARACTERISTICS:
75           (A) LENGTH: 112 amino acids
76           (B) TYPE: amino acid
77           (D) TOPOLOGY: linear
78      (ii) MOLECULE TYPE: protein
79      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
80        Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys
81         1               5                  10                  15
82        Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp
83                       20                  25                  30
84        Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly Pro Cys
85                 35                  40                  45
86        Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu Ala Leu
87            50                  55                  60
88        Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys Val Pro
89         65                  70                  75                  80
90        Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg Lys Pro
91                       85                  90                  95
92        Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys Cys Ser
93                 100                 105                 110
95 (2) INFORMATION FOR SEQ ID NO: 3:
```

RAW SEQUENCE LISTING   DATE: 02/12/2002
PATENT APPLICATION: US/09/813,271B   TIME: 08:59:40

Input Set : N:\Crf3\02062002\I813271B.raw
Output Set: N:\CRF3\02122002\I813271B.raw

```
 96      (i) SEQUENCE CHARACTERISTICS:
 97          (A) LENGTH: 339 base pairs
 98          (B) TYPE: nucleic acid
 99          (C) STRANDEDNESS: double
100          (D) TOPOLOGY: linear
101     (ii) MOLECULE TYPE: cDNA to mRNA
102    (vii) IMMEDIATE SOURCE:
103          (B) CLONE: E. coli LC137/pPLMu.hTGF-beta2 (DSM5657)
104     (ix) FEATURE:
105          (A) NAME/KEY: CDS
106          (B) LOCATION:1..336
107          (D) OTHER INFORMATION:/product= "human TGF-beta2"
108     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
109          GCT TTG GAT GCG GCC TAT TGC TTT AGA AAT GTG CAG GAT AAT TGC TGC      48
110          Ala Leu Asp Ala Ala Tyr Cys Phe Arg Asn Val Gln Asp Asn Cys Cys
W--> 111              115                 120                 125
112          CTA CGT CCA CTT TAC ATT GAT TTC AAG AGG GAT CTA GGG TGG AAA TGG      96
113          Leu Arg Pro Leu Tyr Ile Asp Phe Lys Arg Asp Leu Gly Trp Lys Trp
W--> 114              130                 135                 140
115          ATA CAC GAA CCC AAA GGG TAC AAT GCC AAC TTC TGT GCT GGA GCA TGC     144
116          Ile His Glu Pro Lys Gly Tyr Asn Ala Asn Phe Cys Ala Gly Ala Cys
W--> 117      145                 150                 155                 160
118          CCG TAT TTA TGG AGT TCA GAC ACT CAG CAC AGC AGG GTC CTG AGC TTA     192
119          Pro Tyr Leu Trp Ser Ser Asp Thr Gln His Ser Arg Val Leu Ser Leu
W--> 120                      165                 170                 175
121          TAT AAT ACC ATA AAT CCA GAA GCA TCT GCT TCT CCT TGC TGC GTG TCC     240
122          Tyr Asn Thr Ile Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys Val Ser
W--> 123                      180                 185                 190
124          CAA GAT TTA GAA CCT CTA ACC ATT CTC TAC TAC ATT GGC AAA ACA CCC     288
125          Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Ile Gly Lys Thr Pro
W--> 126              195                 200                 205
127          AAG ATT GAA CAG CTT TCT AAT ATG ATT GTA AAG TCT TGC AAA TGC AGC     336
128          Lys Ile Glu Gln Leu Ser Asn Met Ile Val Lys Ser Cys Lys Cys Ser
W--> 129          210                 215                 220
130          TAA                                                                 339
132 (2) INFORMATION FOR SEQ ID NO: 4:
133     (i) SEQUENCE CHARACTERISTICS:
134         (A) LENGTH: 112 amino acids
135         (B) TYPE: amino acid
136         (D) TOPOLOGY: linear
137     (ii) MOLECULE TYPE: protein
138     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
139         Ala Leu Asp Ala Ala Tyr Cys Phe Arg Asn Val Gln Asp Asn Cys Cys
140          1               5                  10                  15
141         Leu Arg Pro Leu Tyr Ile Asp Phe Lys Arg Asp Leu Gly Trp Lys Trp
142                  20                  25                  30
143         Ile His Glu Pro Lys Gly Tyr Asn Ala Asn Phe Cys Ala Gly Ala Cys
144              35                  40                  45
145         Pro Tyr Leu Trp Ser Ser Asp Thr Gln His Ser Arg Val Leu Ser Leu
```

```
146          50                    55                    60
147      Tyr Asn Thr Ile Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys Val Ser
148          65                    70                    75                    80
149      Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Ile Gly Lys Thr Pro
150                            85                    90                    95
151      Lys Ile Glu Gln Leu Ser Asn Met Ile Val Lys Ser Cys Lys Cys Ser
152                           100                   105                   110
154 (2) INFORMATION FOR SEQ ID NO: 5:
155      (i) SEQUENCE CHARACTERISTICS:
156          (A) LENGTH: 339 base pairs
157          (B) TYPE: nucleic acid
158          (C) STRANDEDNESS: double
159          (D) TOPOLOGY: linear
160     (ii) MOLECULE TYPE: cDNA to mRNA
161    (vii) IMMEDIATE SOURCE:
162          (B) CLONE: E. coli LC137/pPLMu.hTGF-beta3 (DSM 5658)
163     (ix) FEATURE:
164          (A) NAME/KEY: CDS
165          (B) LOCATION:1..336
166          (D) OTHER INFORMATION:/product= "human TGF-beta3"
167     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
168      GCT TTG GAC ACC AAT TAC TGC TTC CGC AAC TTG GAG GAG AAC TGC TGT        48
169      Ala Leu Asp Thr Asn Tyr Cys Phe Arg Asn Leu Glu Glu Asn Cys Cys
W--> 170          115                   120                   125
171      GTG CGC CCC CTC TAC ATT GAC TTC CGA CAG GAT CTG GGC TGG AAG TGG        96
172      Val Arg Pro Leu Tyr Ile Asp Phe Arg Gln Asp Leu Gly Trp Lys Trp
W--> 173          130                   135                   140
174      GTC CAT GAA CCT AAG GGC TAC TAT GCC AAC TTC TGC TCA GGC CCT TGC       144
175      Val His Glu Pro Lys Gly Tyr Tyr Ala Asn Phe Cys Ser Gly Pro Cys
W--> 176      145                   150                   155                   160
177      CCA TAC CTC CGC AGT GCA GAC ACA ACC CAC AGC ACG GTG CTG GGA CTG       192
178      Pro Tyr Leu Arg Ser Ala Asp Thr Thr His Ser Thr Val Leu Gly Leu
W--> 179                   165                   170                   175
180      TAC AAC ACT CTG AAC CCT GAA GCA TCT GCC TCG CCT TGC TGC GTG CCC       240
181      Tyr Asn Thr Leu Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys Val Pro
W--> 182              180                   185                   190
183      CAG GAC CTG GAG CCC CTG ACC ATC CTG TAC TAT GTT GGG AGG ACC CCC       288
184      Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Val Gly Arg Thr Pro
W--> 185          195                   200                   205
186      AAA GTG GAG CAG CTC TCC AAC ATG GTG GTG AAG TCT TGT AAA TGT AGC       336
187      Lys Val Glu Gln Leu Ser Asn Met Val Val Lys Ser Cys Lys Cys Ser
W--> 188      210                   215                   220
189      TGA                                                                  339
191 (2) INFORMATION FOR SEQ ID NO: 6:
192      (i) SEQUENCE CHARACTERISTICS:
193          (A) LENGTH: 112 amino acids
194          (B) TYPE: amino acid
195          (D) TOPOLOGY: linear
196     (ii) MOLECULE TYPE: protein
```

```
         RAW SEQUENCE LISTING              DATE: 02/12/2002
         PATENT APPLICATION:  US/09/813,271B      TIME: 08:59:40

         Input Set : N:\Crf3\02062002\I813271B.raw
         Output Set: N:\CRF3\02122002\I813271B.raw

197     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
198     Ala Leu Asp Thr Asn Tyr Cys Phe Arg Asn Leu Glu Glu Asn Cys Cys
199      1               5                   10                  15
200     Val Arg Pro Leu Tyr Ile Asp Phe Arg Gln Asp Leu Gly Trp Lys Trp
201                     20                  25                  30
202     Val His Glu Pro Lys Gly Tyr Tyr Ala Asn Phe Cys Ser Gly Pro Cys
203                 35                  40                  45
204     Pro Tyr Leu Arg Ser Ala Asp Thr Thr His Ser Thr Val Leu Gly Leu
205             50                  55                  60
206     Tyr Asn Thr Leu Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys Val Pro
207      65                  70                  75                  80
208     Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Val Gly Arg Thr Pro
209                     85                  90                  95
210     Lys Val Glu Gln Leu Ser Asn Met Val Val Lys Ser Cys Lys Cys Ser
211                     100                 105                 110
213 (2) INFORMATION FOR SEQ ID NO: 7:
214     (i) SEQUENCE CHARACTERISTICS:
215         (A) LENGTH: 336 base pairs
216         (B) TYPE: nucleic acid
217         (C) STRANDEDNESS: double
218         (D) TOPOLOGY: linear
219     (ii) MOLECULE TYPE: other nucleic acid
220         (A) DESCRIPTION: /desc = "recombinant hybrid DNA of
221     (vii) IMMEDIATE SOURCE:
222         (B) CLONE: E. coli LC137/pPLMu.TGF-beta1(44/45)beta3
223     (ix) FEATURE:
224         (A) NAME/KEY: mat_peptide
225         (B) LOCATION:1..132
226         (D) OTHER INFORMATION:/product= "N-terminal 44 amino
227 acids of human TGF-beta1"
228     (ix) FEATURE:
229         (A) NAME/KEY: mat_peptide
230         (B) LOCATION:133..336
231         (D) OTHER INFORMATION:/product= "C-terminal 68 amino
232 acids of human TGF-beta3"
233     (ix) FEATURE:
234         (A) NAME/KEY: CDS
235         (B) LOCATION:1..336
236         (D) OTHER INFORMATION:/product= "hybrid TGF-beta named
237 TGF-beta1-3"
238     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
239     GCC CTG GAC ACC AAC TAT TGC TTC AGC TCC ACG GAG AAG AAC TGC TGC    48
240     Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys
241      1               5                   10                  15
242     GTG CGG CAG CTG TAC ATT GAC TTC CGC AAG GAC CTC GGC TGG AAG TGG    96
243     Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp
244                     20                  25                  30
245     ATC CAC GAG CCC AAG GGC TAC CAT GCC AAC TTC TGC TCA GGC CCT TGC   144
246     Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Ser Gly Pro Cys
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/813,271B

DATE: 02/12/2002
TIME: 08:59:41

Input Set : **N:\Crf3\02062002\I813271B.raw**
Output Set: **N:\CRF3\02122002\I813271B.raw**

```
L:21 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:22 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:111 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:114 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:117 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:120 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:123 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:126 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:129 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:170 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:173 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:176 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:179 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:182 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:185 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:188 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
```